The recent evolutionary rescue of a staple crop depended on over half a century of global germplasm exchange

Kebede T. Muleta[1*], Terry Felderhoff[1*], Noah Winans[1], Rachel Walstead[2], Jean Rigaud Charles[3], J. Scott Armstrong[4], Sujan Mamidi[2], Chris Plott[2], John P. Vogel[5], Peggy G. Lemaux[6], Todd C. Mockler[7], Jane Grimwood[2], Jeremy Schmutz[2,5], Gael Pressoir[3], Geoffrey P. Morris[1,8]

[1] Department of Agronomy, Kansas State University, Manhattan, Kansas, USA 66502
[2] Genome Sequencing Center, HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806
[3] Chibas and Faculty of Agriculture and Environmental Sciences, Quisqueya University, Port-au-Prince, Haiti
[4] U.S. Department of Agriculture, Agricultural Research Service, Wheat, Peanut and Other Field Crops Research Unit, 1301 North Western Rd., Stillwater, OK 74075
[5] Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
[6] Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720
[7] Donald Danforth Plant Science Center, St. Louis, MO 63132
[8] Department of Soil & Crop Science, Colorado State University, Fort Collins, Colorado, 80526
[*] Equal contributions

Corresponding author:
Geoffrey P. Morris
Geoff.Morris@colostate.edu

ABSTRACT

Rapid environmental change can lead to extinction of populations or evolutionary rescue via genetic adaptation. In the past several years, smallholder and commercial cultivation of sorghum (*Sorghum bicolor*), a global cereal and forage crop, has been threatened by a global outbreak of an aggressive new biotype of sugarcane aphid (SCA; *Melanaphis sacchari*). Here we characterized genomic signatures of adaptation in a Haitian sorghum breeding population, which had been recently founded from admixed global germplasm, extensively intercrossed, and subjected to intense selection under SCA infestation. We conducted evolutionary population genomics analyses of 296 post-selection Haitian lines compared to 767 global accessions at 159,683 single nucleotide polymorphisms. Despite intense selection, the Haitian population retains high nucleotide diversity through much of the genome due to diverse founders and an intercrossing strategy. A genome-wide fixation (F_{ST}) scan and geographic analyses suggests that adaptation to SCA in Haiti is conferred by a globally-rare East African allele of *RMES1*, which has also spread to other breeding programs in Africa, Asia, and the Americas. *De novo* genome sequencing data for SCA resistant and susceptible lines revealed putative causative variants at *RMES1*. Convenient low-cost markers were developed from the *RMES1* selective sweep and successfully predicted resistance in independent U.S. × African breeding lines and eight U.S. commercial and public breeding programs, demonstrating the global relevance of the findings. Together, the findings highlight the potential of evolutionary genomics to develop adaptive trait breeding technology and the value of global germplasm exchange to facilitate evolutionary rescue.

INTRODUCTION

Ongoing processes of global change, encompassing climate change, nutrient cycles, and pest outbreaks, are shaping the evolution of natural and agricultural ecosystems (*1*, *2*). Intense selection pressure following environment changes may lead to the rapid decline or extinction of populations (*3*, *4*). If a population is to persist under such strong selection, adaptive standing genetic variation must exist or adaptive *de novo* variation must arise on a sufficiently fast timescale (*5*). This population genetic phenomenon, evolutionary rescue, has become a focus of considerable empirical and theoretical study in ecology and conservation biology, since the current rate of global change could exceed the capacities of many populations to adapt (*6*, *7*). Still, there is a lack of examples of evolutionary rescue occuring in the field and at large geographic scales (*4*). In agricultural systems, the spread of pests or emergence of new aggressive biotypes may lead to a reduction of crop diversity or a total loss of crop cultivation (*8*). Therefore, understanding and facilitating evolutionary rescue in agricultural systems is critical for global food security.

Populations of crops or wild species subjected to strong selection pressure may experience a major population bottleneck, resulting in a loss of genetic diversity (*9*). The level of diversity preserved in a population recovering from strong selection depends on the number of backgrounds on which the adaptive alleles emerge (*10*), which can determine the potential for future adaptation or genetic gain. Conversely, adaptation conferred by a beneficial variant derived from a single progenitor causes the removal of genetic diversity from the surviving population (*10*, *11*). Evolutionary population genomics approaches using genome-wide polymorphism data from diverse germplasm can identify candidate loci for adaptive traits (*12*). While genome scans for selection have been widely used to identify putative adaptive alleles in crops (*9*, *13*), they have not yet been used to identify trait-predictive markers for molecular breeding of stress-resilient varieties (*14*).

Sorghum (*Sorghum bicolor* L. [Moench]) is among the world's most important staple crops for smallholder farmers in semiarid regions, as well as a commercial grain and forage crop in industrialized nations (*15*). Since 2013 an aggressive biotype of the sugarcane aphid (SCA; *Melanaphis sacchari*) has become a major threat to global sorghum production, with widespread and substantial yield loss (*16*, *17*). The *M. sacchari* superclone has been rapidly expanding (*18*), putting >90% of the sorghum-producing areas of North America at risk and threatening to end sorghum cultivation in some areas (*16*). In Haiti, a Caribbean nation with one of the world's highest rates of food insecurity, sorghum is among the most important staple crops (*19*). However, heavy infestations by *M. sacchari* since 2015 have caused the loss of over 70% of sorghum production in the country and prevented production of most local landraces (*20*). Shortly before the SCA outbreak, a new Haitian breeding population had been launched by Chibas using global admixed germplasm, rapid-cycling intercrossing, and selection under smallholder conditions (i.e. no insecticidal treatment) (*21*). Selecting from a small number of breeding lines that survived SCA infestation, a new SCA resistant sorghum variety, Papèpichon, was developed and distributed nationally (*19*), and intercrossing and advancement of resistant breeding lines has continued.

Here we used a retrospective genomic analysis of the Haitian sorghum breeding population that was subjected to strong selection under SCA infestation, to understand the genetic basis of the evolutionary rescue following the SCA outbreak, as well as the origins of the SCA resistance alleles. We find that the rapid adaptation of the Haitian breeding population to the SCA outbreak was due to selection for a globally-rare Ethiopian allele at the *RMES1* SCA resistance locus, which is shared across programs in Africa, Asia, and the Americas because of >50 years of global germplasm exchange prior to the SCA outbreak. Further, we developed a convenient low-cost molecular marker based on the evolutionary genome scan and validated it in eight commercial and public sorghum breeding programs, demonstrating the value of leveraging global germplasm exchange and evolutionary population genomics to improve crop resilience.

RESULTS

Genome-wide polymorphism and nucleotide diversity

To understand the evolutionary rescue of sorghum following the SCA outbreak, we conducted a retrospective genomic analysis of the Haitian breeding population (HBP) in comparison to a global diversity panel (GDP). Genotyping-by-sequencing of 296 HBP and 767 GDP (Supp. Fig. S1; Supp. File S1) sorghum lines generated 159,683 polymorphic SNPs with an average SNP density of 75 and 229 per Mb in the HBP and GDP, respectively (Supp. Fig. S2). The GDP had a higher proportion of low-frequency minor alleles (<5% MAF) compared to the HBP (Supp. Fig. S3). Average inbreeding coefficients (F_{IS}) in HBP and the GDP was estimated at 0.7 and 0.9, respectively (Supp. Table S1). The effect of selection on genetic diversity in HBP was assessed based on genome-wide nucleotide diversity (π) in the HBP in comparison to (i) the GDP and (ii) a major public program in the US (Texas A&M pre-breeding lines, TAM-PBL, $N = 35$). Average nucleotide diversity in the HBP was estimated at 2.3×10^{-5}. In the GDP and TAM-PBL, estimates of average π were 5.8×10^{-5} and 4.8×10^{-5}, respectively (Fig. 1A-C, Supp. Table S2). In the HBP, 31% of 1 Mb windows have negative average Tajima's D values, while in the GDP predominantly positive values of Tajima's D were observed (Supp. Fig. S4).

Contributions of global sorghum diversity to the Haitian breeding population

The genetic ancestry of the HBP from global germplasm was inferred based on population structure analyses. In a neighbor joining analysis, the HBP clusters with caudatum accessions (Fig. 2A), specifically caudatums from East Africa. Similarly, in principal coordinate analysis, the HBP cluster swith East African caudatum accessions (Fig. 2C). To estimate ancestry coefficients for HBP lines, we used Bayesian model-based clustering in ADMIXTURE, projecting HBP lines onto ancestral populations and allele frequencies defined using only GDP (with HBP lines omitted). With the GDP, the lowest cross-validation error was observed at $K = 8$ (Supp. Fig. S5) and accessions clustered by ecogeographic region and botanical type, as expected. ADMIXTURE projection analysis suggests that the HBP is admixed, largely consisting of caudatum haplotypes (>80% of the genome) with a remaining small percentage being contributed by durra and guinea sorghums (Fig. 2D).

Evidence of a selective sweep in the Haitian breeding population at *RMES1*

To identify genome regions implicated in the evolutionary rescue of the HBP, genome-wide scans for outlier loci were performed based on an F_{ST} test. Overall, the HBP is moderately differentiated from the global diversity panel, with an average genome-wide F_{ST} of 0.16 (Fig. 3A, Supp. File S2). Based on a Bonferroni-adjusted *P*-value < 0.01, F_{ST} analysis identified 171 outlier genomic regions, which are candidate selective sweep regions. Several genomic regions with F_{ST} outlier regions co-localized with candidate genes for traits under selection by the Chibas breeding program, including photoperiodic flowering, inflorescence architecture, stay□green, stem sugar content, and SCA resistance (Fig 1D; Supp. File S3). Interestingly, the most extreme F_{ST} outliers were observed on chromosome 6, precisely colocalizing with *RMES1*, a locus previously shown to underlie SCA resistance in a Chinese sorghum line of unknown pedigree (*22*) (Fig. 3A-B). To characterize the prevalence of the putative selected haplotype and identify its geographic origin, we mapped the allelic distribution of the highest F_{ST} SNP S6_2995581 in global georeferenced sorghum landraces (Fig. 3C) and compared these distributions to the allele frequency in US and Haitian breeding germplasm (Fig. 3C, inset left). Globally, the allele is rare (<2%), found only in Ethiopian caudatum landraces and a few breeding lines from West Africa and the US. However, the sweep-associated allele is common (~40%) in Ethiopian caudatum accessions (Fig. 3C; Supp. Table S3). The high local frequency of the sweep-associated allele in Ethiopia suggests a likely origin of the SCA resistance allele in the Ethiopian highlands (Fig. 3C, inset right).

Comparative genomic analysis to identify candidate causative variants

To identify candidate causative variants for the *RMES1* locus, we used whole-genome resequencing and *de novo* genome sequencing of sorghum accessions with known SCA reactions. The *RMES1* interval previously defined based on biparental linkage mapping (*22*) includes seven gene models (Sobic.006G017000, Sobic.006G017100, Sobic.006G017200, Sobic.006G017332, Sobic.006G017266, Sobic.006G017400, and Sobic.006G017500) that were candidates for the causative gene. Comparative genomic analyses based on local multiple sequence alignment (MSA) of *de novo* genome sequence of the resistant accession (PI 276837, the Ethiopian progenitor of SCA resistant line SC170) and three sorghum reference genomes of SCA susceptible lines (BTx623, Tx430, and BTx642) were used to identify potential causative variants. No sequence variants were identified in the exons of three of the seven genes (Sobic.006G017000, Sobic.006G017100, and Sobic.006G017266). A total of 35, 32, and 29 nonsynonymous SNPs were detected in the exons of Sobic.006G017200, Sobic.006G017400, and Sobic.006G017500, when comparing the sequences of the resistant PI 276837 and the three susceptible accessions. In addition, three insertion-deletion variations resulting in frame-shift were detected in Sobic.006G017500. (Supp. File S4). To further refine the set of candidate causative variants, we performed a localized association analysis for SCA resistance ("resistant" or "susceptible", based on literature classification) around *RMES1* with resequencing data for diverse sorghum accessions (Fig. 4, Supp. File S5) that detected 101 highly significant associations (*P*-value > 0.0001). Annotations of the variants within the *RMES1* locus indicate that only ten of 101 associated variants are nonsynonymous (5 of 10 in Sobic.006G017200 and the remaining 5 of 10 in Sobic.006G017500.

Development and validation of a molecular marker based on the selective sweep

Next, we sought to test the hypothesis that the genome region identified by the F_{ST} scan underlies variation for SCA resistance in other global sorghum germplasm. Therefore we developed a kompetitive allele specific PCR (KASP) marker based on the SNPs at the *RMES1* locus identified in the F_{ST} scan. Of the candidate SNPs (Supp. File S2, Supp. Table S3, Supp. File S6), SNP 06_02892438 was determined to have the best combination of linkage, LD, and technical KASP functionality of the SNPs. Alternative SNPs were also developed into markers (Supp. File S6), and while the markers are often used as technical checks, testing has confirmed the priority of the marker based on SNP 06_02892438 (Sbv3.1_06_02892438R). Initial validation of the Sbv3.1_06_02892438R KASP marker using DNA samples from known resistant lines (SC110, Tx2783, and IRAT204), susceptible lines (BTx623 and BTx642) (*23*), and multiple F_2 families segregating for SCA resistance demonstrated that the KASP marker Sbv3.1_06_02892438R was in complete agreement with historical phenotypes of inbred lines and segregated within F_2 populations (Supp. File S7).

An F_4 population derived from a cross between IRAT204 (resistant African variety) and Tx430 (susceptible US breeding line) was used to further validate the broader utility and predictiveness of the KASP marker for marker-assisted selection (Fig. 5A-B). A total of 50 F_4 lines together with resistant (IRAT204 and SC110) and susceptible controls (RTx430) were genotyped with the KASP marker Sbv3.1_06_02892438R. Both resistant controls and 23 F_4 lines were homozygous for the resistant allele. The susceptible control and 9 F_4 lines were homozygous for the susceptible allele, and the remaining 18 F_4 lines were heterozygous at the SNP. Twenty-three selected F_4 lines with three resistant and three susceptible control lines were tested for SCA reaction in a free-choice flat screen assay in the greenhouse, scoring aphid damage rating, leaf greenness (SPAD), and seedling height. The SCA reaction phenotypes match the KASP marker genotypes, demonstrating the reliability and predictability of using KASP markers in marker-assisted selection for SCA resistance breeding (Fig. 5A-B; Supp. File S7).

Multi-program validation and deployment in commercial and public breeding programs

To further validate the utility of the SCA resistance KASP markers, we tested them with five US commercial seed company breeding programs and three US public sector breeding programs, representing a large fraction of the US sorghum breeding community (Fig. 5C). (The programs are anonymized to avoid disclosing proprietary information.) Under the hypothesis that (i) *RMES1* underlies SCA resistance in US breeding programs and (ii) the KASP marker (Sbv3.1_06_02892438R) tags the relevant resistant vs. susceptible haplotypes, the breeders' phenotype-based classification of SCA resistance should largely match the KASP marker genotype-based prediction. As expected, the match between the phenotype-based breeder classification and KASP marker genotypes is high, ranging from ~60-100%, with most germplasm sets (9/12) have >80% matching (Fig. 5C; Fig. S6). Less than 0.5% of mismatches (5/1100) were observed among technical replicates (independent tissue samples from the same plant), so mismatches are unlikely to be due to KASP genotyping errors. Mismatches may be due to differences among programs of SCA resistant or susceptible haplotypes, or errors in the phenotype-based resistance classifications (some of which are based on visual ratings under natural field infestations, which are prone to false positives (*24*)). There were also some genotype-phenotype mismatches in public germplasm checks used by commercial and public programs (Fig. 5C). In nearly all cases, further investigation revealed that mismatches were due to unexpected heterogeneity in public germplasm within or among breeding programs (Supp. Table S4).

DISCUSSION

RMES1 is a major resistance gene underlying evolutionary rescue of sorghum worldwide

Understanding the genetics of evolutionary rescue, including the genetic architecture and molecular basis, could contribute to more resilient conservation and breeding strategies (*25*). Here we hypothesized, parsimoniously, that a single Mendelian SCA resistance locus *RMES1* could underlie the global evolutionary rescue of sorghum to the new *M. sacchari* superclone. Previous studies had suggested that a single dominant locus is responsible for SCA resistance in families derived from resistant Chinese grain sorghum variety Henong 16 (H16) and susceptible BTx623, or families derived from US breeding lines, resistant RTx2738 and susceptible CK60 (*22*, *26*). The H16 resistance was mapped to a ~130 kb region at 2.7 Mb on chromosome 6 (*RMES1*) (*22*). Consistent with the *RMES1* evolutionary rescue hypothesis, the genome region with the highest F_{ST} in the HBP colocalized precisely with *RMES1* (Fig. 3). Together, the evolutionary genome scan (Fig. 3) and multi-program marker validation (Fig. 5) provides strong evidence that *RMES1* is the major SCA resistance locus globally, shared across the Americas, Asia, and Africa. However, our findings do not preclude the hypothesis that other SCA resistance loci were selected in Haiti and were required for the evolutionary rescue. In particular, other F_{ST} scan peaks on chromosome 2, 7, 8, and 9 (Fig. 3) could correspond to other SCA resistance loci. Given that SCA resistance is fixed in the Haitian program, further population development and quantitative trait locus mapping for SCA resistance will be necessary to test this hypothesis.

Identifying the causal variant underlying SCA resistance would advance our understanding of aphid resistance mechanisms in plants (*27*) and facilitate development of perfectly-predictive molecular markers for SCA resistance breeding (*28*). Our comparative genomic analysis between the resistant PI 276837 and the three susceptible reference genomes identified four candidate genes with putative functional variants within the *RMES1* locus (Sobic.006G017200, Sobic.006G017332, Sobic.006G017400 and Sobic.006G017500; Supp. File S4). Three of the four genes in the candidate region encode leucine-rich repeat (LRR) proteins, a gene family involved in immune responses to invading pathogens and insects (*29*). Given that some LRR genes mediate plant resistance to aphids and other phloem-feeding insects (*27*) these genes represent promising candidates for the *RMES1* causative gene. Functional annotation and sequence comparison between the resistant and susceptible accession identified non-synonymous variants only in Sobic.006G017200 and Sobic.006G017500 (Fig. 4), suggesting these NLR are promising candidates for the *RMES1* gene. Fine-mapping and positional cloning will be needed to test these hypotheses and positively identify the causative variant.

Evolutionary rescue of sorghum depended on a half century of global germplasm exchange

In the twentieth century, sorghum genebanks and breeding programs exchanged germplasm widely (*30*, *31*). Based on pedigree records and morphology we hypothesized that the Haitian breeding population originated from global admixed germplasm with a primary contribution of Ethiopian caudatum of the zerazera working group. Consistent with this hypothesis, HBP genotypes clustered with caudatum sorghum of East Africa (Fig. 2), but admixture analysis identified a contribution from durra and guinea sorghum from West Africa (Fig. 2D). Combining population genomics findings (Fig. 2, 3) with genebank and pedigree records (*32*, *33*), we can map the history of global germplasm exchange that led to the evolutionary rescue of sorghum in Haiti following the SCA outbreak (Fig. 6A), as well as the spread of the SCA resistance allele from Ethiopia to breeding programs around the world (Fig. 6B). Notably, the evolutionary rescue of sorghum in the Americas (Haiti and US) involved germplasm and knowledge exchange over a period of >50 years, involving nine countries on three continents.

In the case of the SCA outbreak, the global sorghum improvement community was fortunate that the rare SCA resistance allele originated in East African caudatums, since this germplasm is preferred by many sorghum breeders worldwide and widely used by breeding programs in Africa, Asia, and the Americas (*30*, *34*). The SCA resistance allele appears to have been inadvertently spread across sorghum breeding programs across the three continents long before the recent SCA outbreak (Fig. 6). For example, SC110, a converted version of an Ethiopia caudatum landrace (PI 257599/IS 12610) identified as SCA resistant in several world regions (*23*, *35*), is a major contributor to the pedigrees of most SCA-resistant breeding lines in the US (Fig. 6B) (*36*). The same progenitor line (IS 12610) was used by breeding programs in West Africa (Fig. 6B) as a parent of IRAT204 (CE151-262; PI 656031), a widely-adopted variety (*37*) and key progenitor of current West African breeding programs (*34*).

Another potential benefit of germplasm exchange is the maintenance of diversity in breeding programs following strong selection, including evolutionary rescue. Given the strong selection on the HBP during the SCA outbreak, it might be expected that the post-selection HBP no longer retains sufficient diversity for future adaptation and genetic gain (*7*). However, the HBP was founded with diverse admixed global germplasm (Fig. 2) and extensively intercrossed, so it appears to have retained sufficient genetic diversity for future adaptation and crop improvement. We observe only a modest reduction in nucleotide diversity observed throughout the genome of the HBP relative to global accessions, East African caudatums, or a major public pre-breeding program (Fig. 1E; Supp. Fig. S7). Recombination during intercrossing cycles (prior to the SCA outbreak) presumably reshuffled the SCA resistance allele onto many backgrounds, suggesting that the intercrossing approach was critical to allow the Haitian program to retain diversity for future genetic gain and adaptation.

Rapid discovery and deployment of a global trait-predictive molecular marker using evolutionary population genomics

Molecular marker development based on phenotype-to-genotype mapping of trait loci (e.g. linkage or association mapping) is limited by availability of suitable mapping populations, phenotyping capacity, and genotyping resources, which can take years to develop (*13*, *38*). For instance, spatial and temporal variability of SCA pressure in field trials limits the effectiveness of field phenotyping (*24*), while greenhouse assays can be complicated and time-consuming for lower-resourced programs. Thus, an evolutionary genomics approach, which leverages a history of selection by smallholder farmers or plant breeders, could have advantages for marker discovery. Despite wide use of evolutionary genome scans in crops, the hypotheses generated on adaptive loci are rarely, if ever, tested by independent experimental approaches (e.g. with near isogenic lines) (*39*). To our knowledge, this is the first example where an evolutionary or population genomic scan led directly to molecular breeding technology in use in commercial and public varietal development (Fig. 3, 5).

Here we demonstrated the effectiveness of the evolutionary population genomic approach, showing that a marker discovered in a single developing-country breeding program (Chibas-Haiti) can link crop improvement efforts across three continents (North America, Africa, Asia; Fig. 5A, 6) and across the commercial and public sector (Fig. 5A, C). Thus, our findings establish the value of evolutionary population genomics to facilitate and guide global crop improvement. The KASP marker developed and validated in this study can facilitate the rapid conversion of existing farmer-preferred varieties for SCA resistance (e.g. by marker-assisted introgression) (*40*). While the *RMES1* resistance allele is currently conferring effective resistance, a further biotype shift in the aphid could overcome this gene. Several biotype shifts occurred in the 1960-1980s for the greenbug aphid *Schizaphis graminum* (*41*) and slowed genetic gain in sorghum for many years (*42*). The markers developed here could facilitate identification of new SCA resistance genes, via by counterselection of *RMES1* allele to reveal novel SCA resistance. These outsourced KASP markers are convenient for breeding programs, since they require no laboratory labor or facilities, and are low cost relative to dedicated field or greenhouse phenotyping capacity, at ~$2 per sample for DNA extraction and marker genotyping (*43*).

Synergy of long-standing germplasm exchange practices with new genomics technologies

In this study, we integrated evolutionary population genomic analyses and historical records on global germplasm exchange to show that the recent evolutionary rescue of sorghum depended on >50 years of germplasm exchange. Germplasm exchange led to global diffusion of a rare SCA resistance allele, sometimes purposely and sometimes inadvertently, from smallholder farmers in Ethiopian highlands across breeding programs in Africa, the Americas, and Asia. Over the past several decades, movement of crop genetic resources through international cooperation of germplasm exchange has provided access to adaptive genetic variation for crop improvement (*31*). However, germplasm exchange is increasingly restricted due to commercial or institutional interests asserting intellectual property (IP) rights and governments asserting national or local sovereignty over genetic resources (*44*). While IP rights and sovereignty are important considerations, the question remains how to balance these aims with the benefits of free exchange of global public goods (*45*, *46*). While we are not in a position to resolve these societal tradeoffs, our study does highlight the global food security benefits of germplasm exchange and the opportunities that could be lost due to restrictions on exchange.

Taken together, our findings suggest that new genomic technologies will be most powerful when leveraged with global exchange of germplasm and knowledge. No matter how powerful new genomic technologies are in terms of accuracy or throughput, their utility will depend on the germplasm assayed, since all genetic mapping approaches require effective recombination and allelic diversity (*47*, *48*). Global germplasm exchange vastly increases both these parameters, providing a "bank" of historical recombinations and allelic variants that can be rapidly leveraged with new genomic tools (Fig. 3, 4). Therefore, our best opportunity to address challenges of global change may be to leverage new genomic technologies with long-standing practices of global germplasm exchange.

MATERIALS AND METHODS

Sorghum breeding and production in Haiti

The Chibas sorghum breeding program was launched in 2013 using admixed global germplasm, including heterogeneous breeding material from West Africa carrying *ms3* nuclear male sterility, and inbred global accessions. During 2015–2018, the material was selected in breeding nurseries under low-input conditions (approximating local smallholder practices) and extensively intercrossed using the *ms3* sterility system. No insecticides were used to limit SCA infestations in breeding nurseries in this period and natural SCA infestations were intense during this period (e.g. Fig. 1C). Note, selection pressure on sorghum by SCA in Haiti is expected to be greater than in temperate zone (e.g. U.S.) because the SCA infestation occurs year-round in this tropical environment. Annual sorghum production estimates for Haiti are based on FAOSTAT (2009–2014 and 2018) (*49*) and the USDA forecast for 2019-2020 (*19*). FAOSTAT data for 2015–2017 and 2019 was not used because it was based on imputation ("FAO data based on imputation methodology") that did not account for the known effects of SCA (e.g. "this aphid spread throughout the country and decimated Haiti sorghum production") (*19*). Production for the missing years of SCA outbreak was inferred based on 2009 agriculture survey acreage prior to infestation in each region and assessment of sorghum production levels compared to pre-infestation levels, adjusted to FAOSTAT (1990–2014) production averages for each region.

Plant genetic resources

The HBP (N = 296) are inbred lines derived from a recurrent selection breeding population developed by intercrossing germplasm that survived natural SCA infestation. For genomic DNA extraction, fresh leaf tissue of each accession was collected from two weeks old seedlings raised in a greenhouse. Tissue was lyophilized for two days and then grounded up using a 96-well plate plant tissue grinder (Retsch Mixer Mill). Genomic DNA was extracted using the BioSprint 96 DNA Plant Kit (QIAGEN), quantified using Quant-iT™ PicoGreen® dsDNA Assay Kit, and normalized to 10 ng/uL. An additional set of global accessions (GDP, $N = 767$) was assembled based on a published data set (*50*, *51*) including sorghum accessions from 52 countries on five continents and all major botanical types (Supp. Fig. S1, Supp. File S1). The GDP accessions included 164 caudatum, 96 guinea, 81 durra, 57 bicolor, and 47 kafir accessions, along with 288 of other botanical types and 34 accessions of unknown botanical type.

Genotyping-by-sequencing

Genotypes for the 296 Haitian breeding lines were generated with genotyping-by-sequencing. Genomic DNA digestion, ligation and PCR amplification processes were performed according to the methods previously described (*50*). The libraries were sequenced using the single-end 100-cycle sequencing by Illumina HiSeq2500 (Illumina, San Diego CA, USA) at the University of Kansas Medical Center, Kansas City, MO, USA. A total of 220 million reads for the HBP were combined with published data for the GDP (*50*) for SNP calling. TASSEL 5 GBS v2 pipeline (*52*) was used to perform the SNP calling of the sequence data obtained from Illumina sequencing. Reads were aligned to the BTx623 sorghum reference genome v.3.1 (*53*) with the Burrows-Wheeler Alignment (*54*). The SNPs were filtered for 20% missingness, then missing data were imputed using BEAGLE 4.0 (*55*). Genotyping data are available at Dryad [*accession to be added following acceptance*].

Population genomic analyses

Genome-wide nucleotide diversity (π) and Tajima's D statistics for HBP and GDP were estimated based on a non-overlapping sliding window of 1 Mbp across the genome using VCFtools (*56*). The characterization of the population structure of the HBP was based on a discriminant analysis of principal components (DAPC) in the Adegenet package in R (*57*). A distance matrix calculated based on a modified Euclidean distance model was used to create a cladogram based on a neighbor-joining algorithm in TASSEL (*58*). Neighbor-joining analysis was visualized using the APE package in R (*59*). The population structure of the germplasm panel was further assessed by the Bayesian model-based clustering method implemented in the ADMIXTURE program (*60*). Pairwise SNP differentiation (F_{ST}) between the HBP and the GDP were calculated and outlier loci were detected based on an inferred distribution of neutral F_{ST} using the R Package OutFLANK (*61*).

Whole genome resequencing

Around the 130 kb mapped interval in BTx623, SNPs from 10 sorghum accessions with known SCA resistance status were examined to search for functional mutations responsible for SCA resistance. Six of the 10 resequenced accessions represent known susceptible lines, which include RTx430 (PI 655996), BTx623 (PI 564163), Tx7000 (PI 655986), Tx2737 (PI 655978), BTx642, and RTx436. The remaining four resequenced accessions represent known resistant lines, which includes PI 257599 (SC110 original exotic parent), PI 276837 (SC170 original exotic parent), PI 534157 (SC170), and IS 36563 (IRAT204). These samples were used pre-publication for this interval analysis with permission from TERRA-REF (Mockler), JGI Sorghum Pan-genome project (Mockler), BMFG Sorghum Genomic Toolbox (Mockler and

Morris), JGI Sorghum Diversity project (John Mullet), and the JGI EPICON project (Vogel). The reads were mapped to *Sorghum bicolor* v3.1 using bwa-mem. The bam file was filtered for duplicates using Picard (http://broadinstitute.github.io/picard) and realigned around indels using GATK (*62*). Multi-sample SNP calling was done using SAMtools mpileup and Varscan V2.4.0 with a minimum coverage of 8 and a minimum alternate allele frequency of four. Repeat content of the genome was masked using 24 bp kmers. Kmers that occur at a high frequency, up to 5%, were masked. SNPs around 25 bp of the mask were removed for further analysis. A SNP was included for further analysis only when it has coverage in 75% of the samples, and a MAF > 0.005. Functional annotation of the variants within the *RMES1* locus was performed using SNPEff.

***De novo* genome sequencing**

De novo genome assembly of the resistance sorghum line PI 276837 was used to perform comparative genomic analysis to identify the causative variant for SCA resistance at the *RMES1* locus. PI 276837 main assembly consisted of 101.47x of PACBIO coverage with an average read size of 11,931 bp. The genome was assembled using Canu 1.8, a fork of the Celera Assembler designed for high-noise single-molecule sequencing. The resulting sequence was polished using ARROW. The assembled genome resulted in contig N50 sizes ranging from 14 to 19 kb and scaffold N50 sizes ranging from 5 to 65 kb. Sequence variations at *RMES1* locus between the *de novo* sequence of PI 276837 were compared to the reference genomes of BTx623, Tx430, and BTx642.

KASP marker development

SNPs from the F_{ST} genomic selection scan were selected for development into markers based on several factors: LOD score of the F_{ST} analysis, proximity to *RMES1* locus, and suitability of the flanking sequence for KASP assay development. The KASP assays were developed utilizing a third-party genotyping service provider, Intertek AgriTech (Alnarp, Sweden), who designed the KASP assays via the Kraken software. All genomic DNA extraction and KASP genotyping were performed by Intertek using two 6 mm leaf punches dried with silica beads. Initial technical validation of the KASP marker was performed using known resistant (SC110, IRAT204 and Tx2783) and susceptible (KS585 and Tx7000) sorghum lines. Further validation was performed by genotyping a panel of 10 known resistant and 28 known susceptible lines, along with multiple F1 crosses of each of the lines. The KASP markers developed for SCA resistance selection (Supp. File S6) are publicly available through the third-party genotyping service provided by Intertek. For further information on accessing markers contact the corresponding author.

Marker validation in public and commercial breeding programs

To test the predictiveness of the marker, a population segregating for SCA resistance was developed by crossing the susceptible Tx430 and resistant IRAT204. F_3 and F_4 lines of the Tx430 x IRAT204 population were genotyped with the KASP marker together with the susceptible and resistant parents. The same population was evaluated for SCA reaction using a free-choice flat-screen trial in the greenhouse. Tx2783 and SC110 were included as known resistant genotypes, along with the known susceptible genotypes, KS 585 and Tx7000 (*63*, *64*). Free-choice flat-screen assay, data collection (damage rating, SPAD score, and plant height difference), and analysis were conducted as previously described (*24*). Validation of the KASP marker across different breeding programs was performed in eight breeding programs, five commercial and three public in the US. Each program collected tissue samples from known tolerant and susceptible parental breeding lines, F_1s of the parental lines, and later generation lines from their SCA tolerance breeding populations; the SCA reaction phenotypes of the late generation lines may or may not have been known. For the parental breeding lines, both technical replicates (tissue samples from the same plant) and biological replicates (tissue samples from separate plants) were collected in order to test both the technical function of the markers and the reliability of the germplasm, respectively. Additionally, most programs included public sources (e.g. Tx2783) of known SCA tolerance as checks. Tissue samples were sent to Intertek, who extracted DNA and performed the KASP genotyping.

ACKNOWLEDGMENTS

This study is made possible by the support of the American People provided to the Feed the Future Innovation Lab for Collaborative Research on Sorghum and Millet through the United States Agency for International Development (USAID) under Associate Award No. AID-OAA-LA-16-00003, “Feed the Future Innovation Lab for Genomics-Assisted Sorghum Breeding”. The contents are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government. The work conducted by the US Department of Energy Joint Genome Institute is supported by the Office of Science of the US Department of Energy under Contract No DE-AC02-05CH11231. We thank the Joint Genome Institute and collaborators for pre-publication access to the genomes of Tx430 and Tx642 for use in this study. Additional support was provided by the Bill & Melinda Gates Foundation under the "Sorghum Genomics Toolkit" project. We thank the breeding programs that participated in the marker testing. We thank Matt Davis for technical support and Jesse Lasky for comments on the manuscript. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable. USDA is an equal opportunity provider and employer. The authors declare that they have no competing interests. All data needed to evaluate the conclusions in the paper are present in the Supplementary Materials or Dryad.

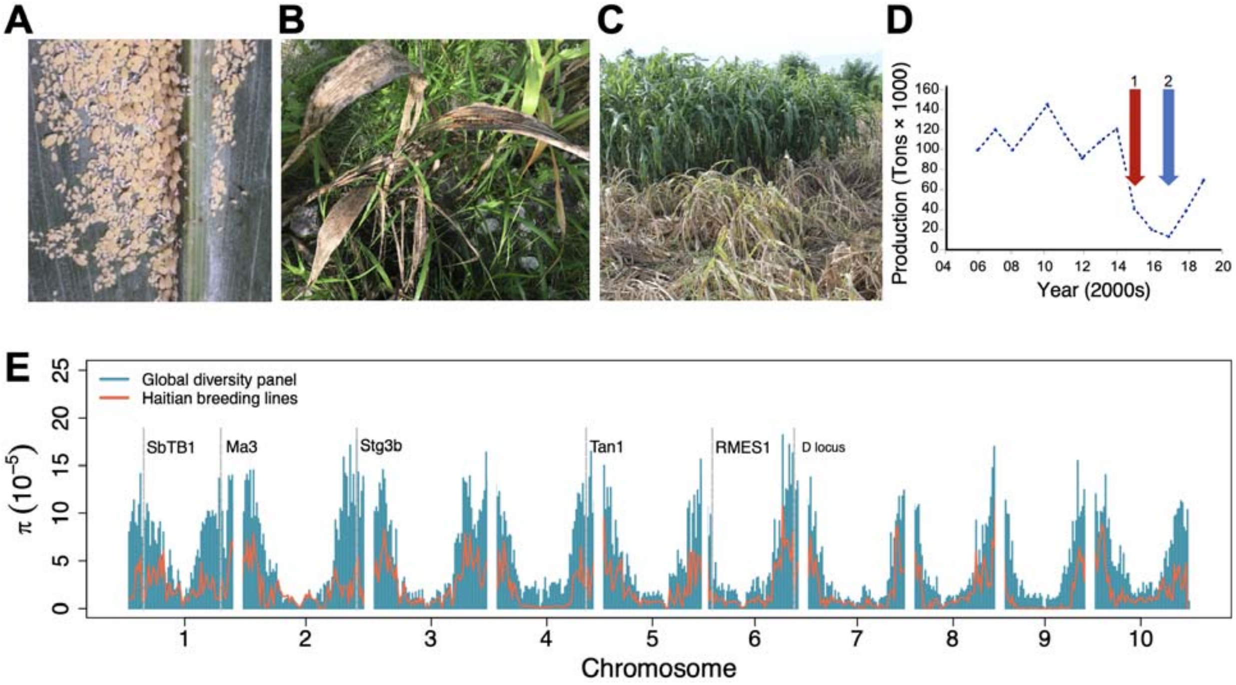

Figure 1: Evolutionary rescue following a continental outbreak of a sorghum pest (A) Infestation of sugarcane aphid (SCA), *M. sacchari*, on a commercial hybrid in the US sorghum-growing production region (Kansas). (B) SCA infestation on a traditional sorghum variety on a smallholder farm in Haiti (brown plant in foreground; green leaves in background are maize and wild grasses). (C) Reaction of susceptible (brown plants; foreground) and resistant (green plants; background) sorghum breeding lines under natural SCA infestation during breeding trials in Haiti. (D) Estimates of annual sorghum production in Haiti (2006-2019), indicating the start of the SCA outbreak (1, red arrow) and the start of national distribution of SCA resistant variety, Papépichon (2, blue arrow). (E) Genome-wide nucleotide diversity (π) in the Haitian breeding population (red line) compared to a global diversity panel (blue bars). Nucleotide diversity was calculated for a non-overlapping sliding window of 1 Mbp across the genome. The grey vertical dashed lines indicate the position of *a priori* candidate genes for breeding targets of the Haiti program which colocalized with genomic regions of reduced π (see Supp. File S3 for details).

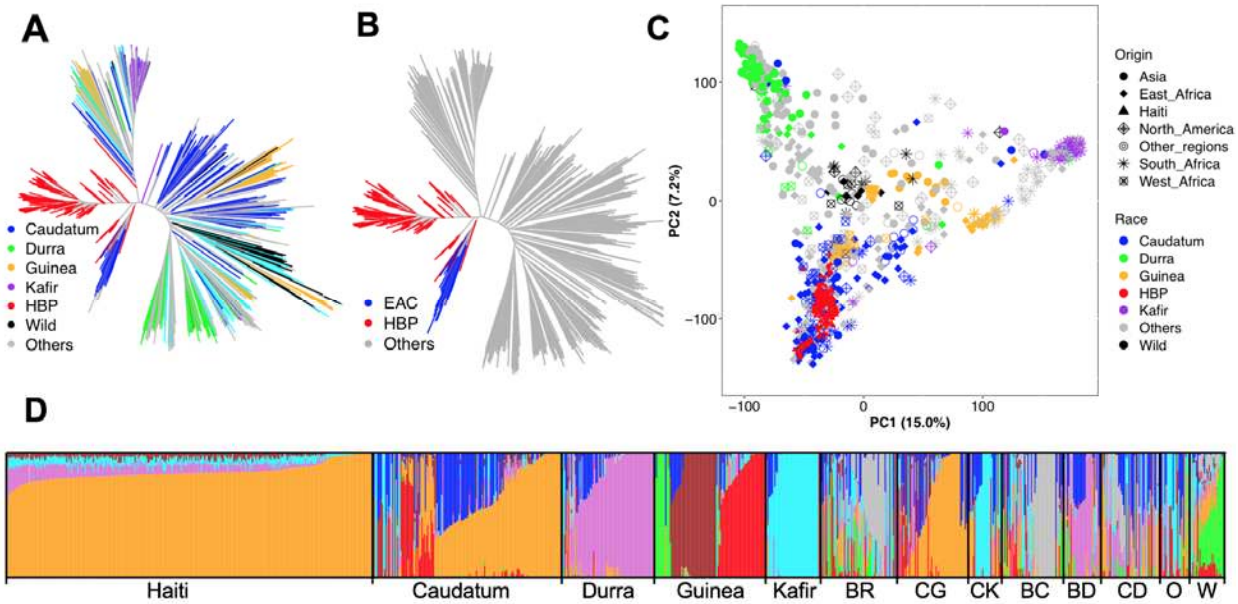

Figure 2: **Population structure of the Haitian breeding population in relation to global sorghum diversity reflects its derivation from East African germplasm**. Genetic relatedness of the Haitian breeding population (HBP) to the global diversity assessed by neighboring joining method, color-coded by botanical type (A) or highlighting the close relationship between the HBP and East African caudatum (EAC) germplasm (C) Scatterplot of the first two principal component (PC) of genome wide SNP variation, demonstrating the clustering of HBP within EAC germplasm. (D) Bayesian hierarchical clustering of the HBP and GDP with the probability of membership (*Q*) in each of K = 8 ancestral populations. The Q-value bar plots are arranged by botanical types to reflect the relationship of the HBP to the GDP. Note, color-coding of the bar plots in panel D is arbitrary and does not reflect the color-code in panels A-C. BR = Bicolor, CG = caudatum-guinea, CK= caudatum-kafir, BC = bicolor-caudatum, BD = bicolor-durra, CD = caudatum durra, O = others (includes botanical types containing less than 10 individuals), W= wild.

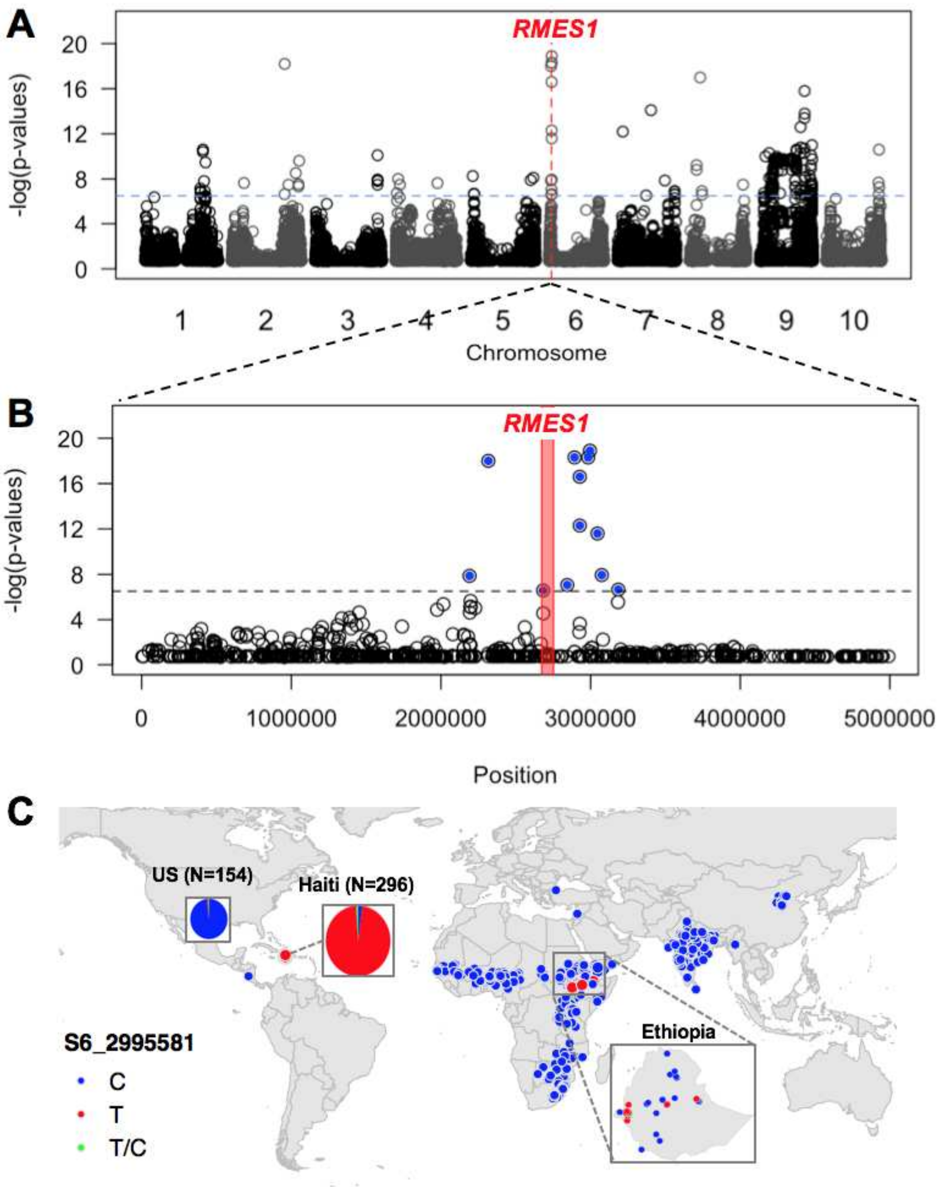

Figure 3: **Genome scan for selection identifies the major aphid resistance allele at *RMES1* originating in Ethiopia.** (A) Genome-wide scan for selection in the Haitian breeding population using fixation index (F_{ST}) with the -log(F_{ST} p-value) (y-axis) plotted against position on the chromosome (x-axis). (B) Detailed view (5 Mb) of top F_{ST} peak on chromosome 6 that colocalizes with the *RMES1* locus. The ~130 kb region from 2,667,082 to 2,796,847 bp corresponding to the published *RMES1* interval is denoted with the red bar. (C) Global allele distribution of the SNP that showed the highest F_{ST} value (S6_2995581), which colocalized with the *RMES1* locus. Allelic state for georeferenced global germplasm is denoted with points. Allele frequencies in the United States (C=151, T=2, T/C=1) and Haiti (C=6, T=287, T/C =3) breeding germplasm, denoted in pie charts with area proportional to number of accessions, show the allele is almost fixed in Haitian breeding germplasm and rare in U.S. breeding germplasm.

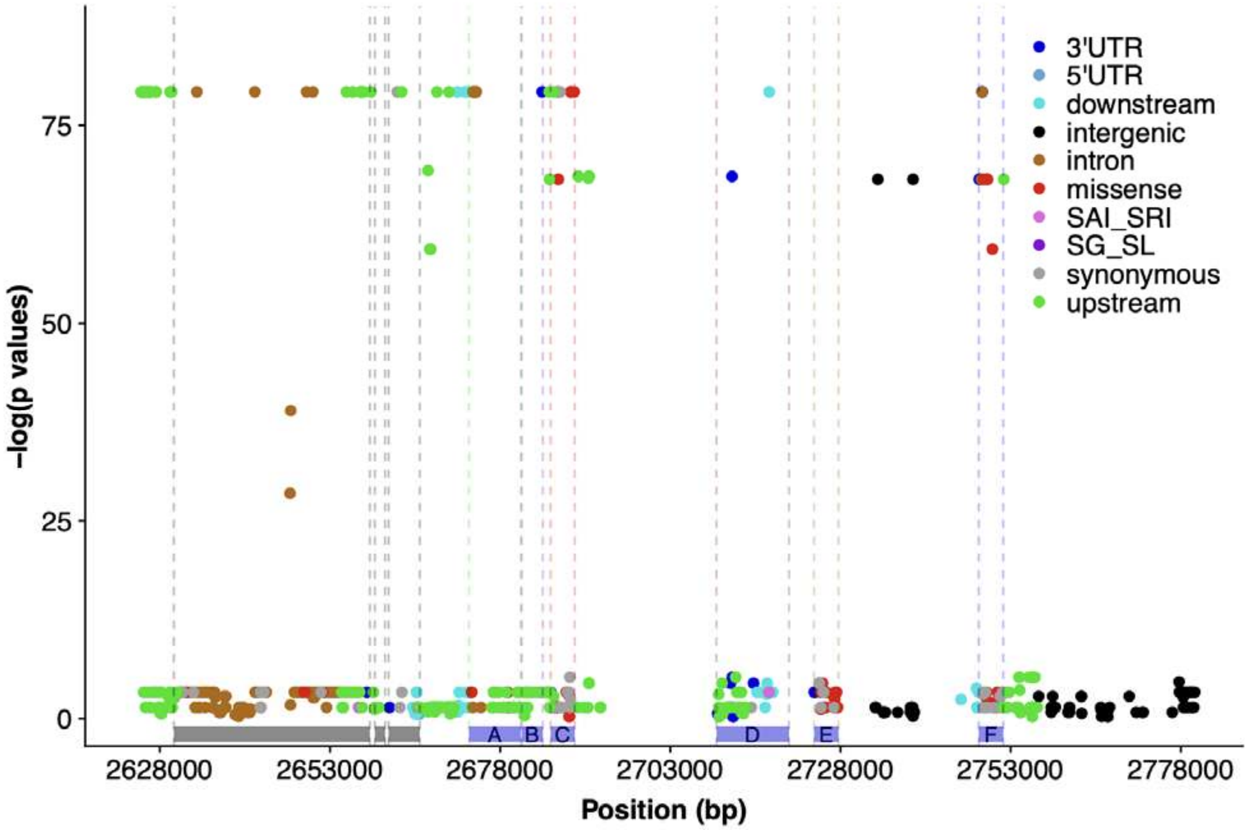

Figure 4: Whole-genome resequencing and local association mapping identifies potential causative variants at *RMES1*. Functional annotation and association mapping of nucleotide polymorphisms within the *RMES1* locus across a set of 13 diverse sorghum accessions with known SCA resistance or susceptibility. The -log of *p*-values of local marker-trait association scan plotted against the chromosomal positions at the *RMES1* locus on chromosome 6. Variants are color-coded by annotation generated by the SnpEff program. Blue bars represent the seven annotated genes within the *RMES1* interval (A = Sobic.006G017000, B = Sobic.006G017100, C = Sobic.006G017200, D = Sobic.006G017332 and Sobic.006G017266, E = Sobic.006G017400, F = Sobic.006G017500.v3.1). Grey bars indicate genes outside of the *RMES1* interval as originally defined (*22*). 3'UTR: 3 prime UTR variant, 5'UTR: 5 prime UTR variant, downstream: Downstream gene variant, intergenic: intergenic region, intron: Intron variant, missense: Missense variant, SAI_SRI: splice acceptor/intron or splice region intron variants, SG_SL: stop gained or stop loss variant, synonymous: Synonymous variant upstream: Upstream gene variant

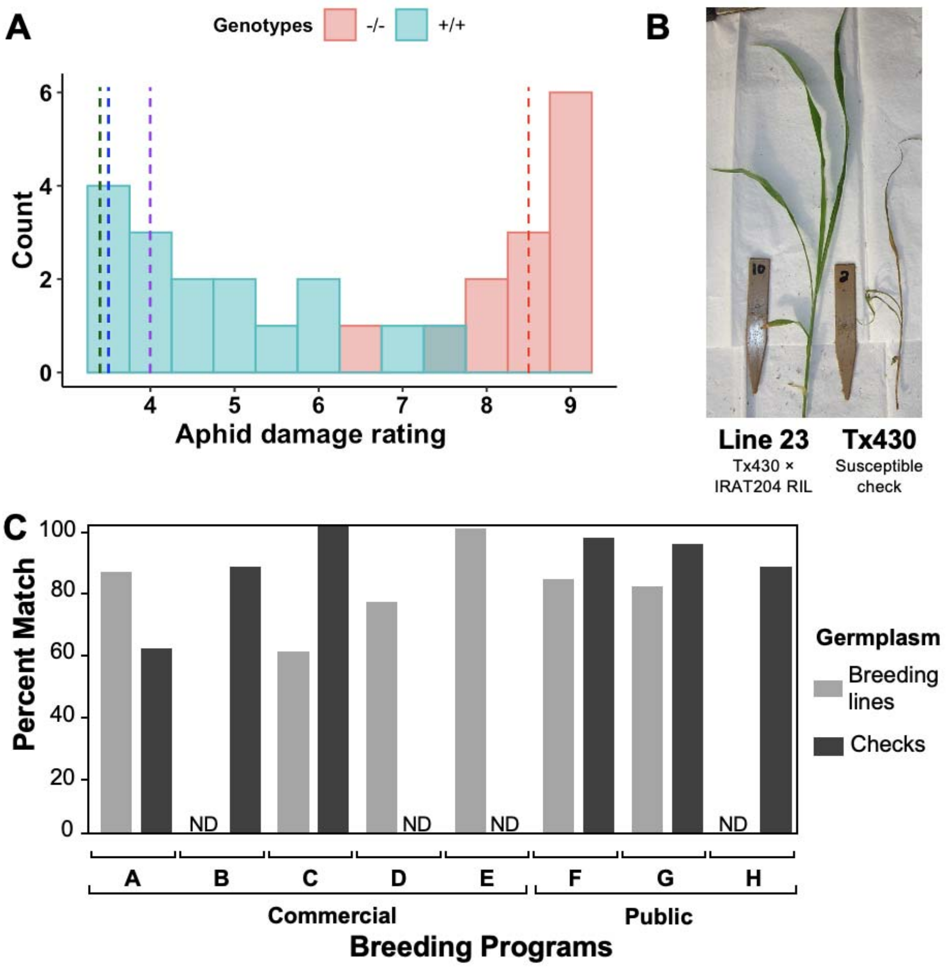

Figure 5: Multi-program evaluation of a molecular marker developed based on the selective sweep validates its global utility.
(A) The KASP marker predicts SCA resistance in independent US × Senegal breeding lines. The histogram represents the aphid damage ratings of F$_4$ lines from a Tx430 × IRAT204 (N = 22) family under infestation by *M. sacchari* at the seedling stage in a choice greenhouse assay. The cyan bars represent the aphid damage ratings for lines carrying the +/+ genotypes at the SNP 06_02892438, while the red bars represent aphid damage ratings of the lines carrying the -/- genotypes. The blue dashed lines represent the average aphid damage rating of the resistant checks Tx2783, IRAT204, and SC110 (green, blue, and purple dashed lines, respectively) while red dashed line represents the average damage rating of both susceptible checks, Tx7000 and Tx430. (B) Representative SCA reaction from the choice greenhouse assay for an F4 line carrying the +/+ genotype (left) versus the susceptible parent Tx430 (right). (C) Evaluation of the same marker in eight US breeding programs. Percent match of KASP marker genotyping prediction with breeder-provided SCA resistance classification for five commercial breeding programs and three public breeding programs. ND = Not determined.

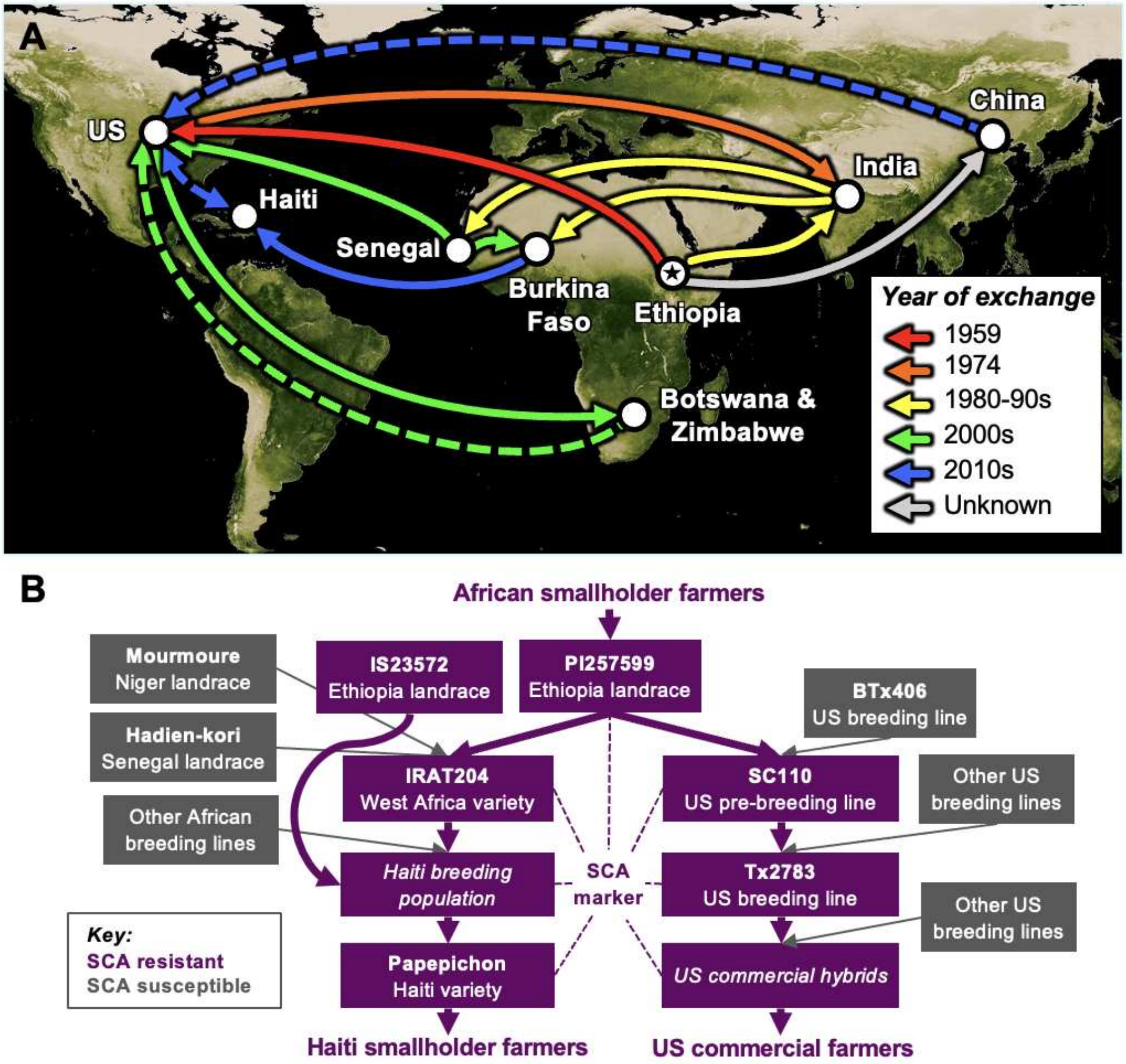

Figure 6: Evolutionary rescue of sorghum through >50 years of global exchange of germplasm and knowledge. (A) Germplasm and knowledge exchange inferred from pedigree records and genomic analyses. Germplasm exchange is denoted by solid lines. Knowledge exchange through scientific literature is denoted in dashed lines. The star indicates the inferred origin of the SCA resistance allele in the Ethiopian highlands, with at least two paths to the Americas, via IS 23572 (yellow line) or PI 257599 (red line). (B) Pedigree relationships among global accessions, breeding lines, breeding programs, or varieties, color-coded by inferred SCA resistance or susceptibility. Note, with respect to US commercial hybrids, the diagram is illustrative and is not meant to imply that all US commercial hybrids used Tx2783 as the SCA resistance donor. Some known pedigree information has been omitted from the diagram for clarity.

REFERENCES

1. S. R. Palumbi, Humans as the world's greatest evolutionary force. *Science*. **293**, 1786–1790 (2001).
2. T. Wheeler, J. von Braun, Climate change impacts on global food security. *Science*. **341**, 508–513 (2013).
3. Gonzalez, Ronce, Ophélie, Ferriere, Regis, Hochberg, Michael E., Evolutionary rescue: an emerging focus at the intersection between ecology and evolution. *Philos. Trans. R. Soc. B Biol. Sci.* **368**, 20120404 (2013).
4. G. Bell, Evolutionary Rescue. *Annu. Rev. Ecol. Evol. Syst.* **48**, 605–627 (2017).
5. B. M. Peter, E. Huerta-Sanchez, R. Nielsen, Distinguishing between Selective Sweeps from Standing Variation and from a De Novo Mutation. *PLOS Genet.* **8**, e1003011 (2012).
6. H. K. Alexander, G. Martin, O. Y. Martin, S. Bonhoeffer, Evolutionary rescue: linking theory for conservation and medicine. *Evol. Appl.* **7**, 1161–1179 (2014).
7. H. A. Orr, R. L. Unckless, The population genetics of evolutionary rescue. *PLOS Genet.* **10**, e1004551 (2014).
8. R. S. Turner, After the famine: Plant pathology, Phytophthora infestans, and the late blight of potatoes, 1845–1960. *Hist. Stud. Phys. Biol. Sci.* **35**, 341–370 (2005).
9. J. M. Burke, J. C. Burger, M. A. Chapman, Crop evolution: from genetics to genomics. *Curr. Opin. Genet. Dev.* **17**, 525–532 (2007).
10. B. A. Wilson, P. S. Pennings, D. A. Petrov, Soft Selective Sweeps in Evolutionary Rescue. *Genetics*. **205**, 1573–1586 (2017).
11. P. S. Pennings, J. Hermisson, Soft Sweeps III: The Signature of Positive Selection from Recurrent Mutation. *PLOS Genet.* **2**, e186 (2006).
12. J. J. Vitti, S. R. Grossman, P. C. Sabeti, Detecting natural selection in genomic data. *Annu. Rev. Genet.* **47**, 97–120 (2013).
13. C. R. Cavanagh, S. Chao, S. Wang, B. E. Huang, S. Stephen, S. Kiani, K. Forrest, C. Saintenac, G. L. Brown-Guedira, A. Akhunova, D. See, G. Bai, M. Pumphrey, L. Tomar, D. Wong, S. Kong, M. Reynolds, M. L. da Silva, H. Bockelman, L. Talbert, J. A. Anderson, S. Dreisigacker, S. Baenziger, A. Carter, V. Korzun, P. L. Morrell, J. Dubcovsky, M. K. Morell, M. E. Sorrells, M. J. Hayden, E. Akhunov, Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl. Acad. Sci.* (2013), doi:10.1073/pnas.1217133110.
14. M. B. Hufford, J. C. Berny Mier y Teran, P. Gepts, Crop Biodiversity: An Unfinished Magnum Opus of Nature. *Annu. Rev. Plant Biol.* **70**, 727–751 (2019).
15. R. Monk, C. Franks, J. Dahlberg, in *Yield Gains in Major US Field Crops* (Crop Science Society of America, 2014; https://dl.sciencesocieties.org/publications/books/abstracts/cssaspecialpubl/yieldgainsinmaj/293), pp. 293–310.
16. R. D. Bowling, M. J. Brewer, D. L. Kerns, J. Gordy, N. Seiter, N. E. Elliott, G. D. Buntin, M. O. Way, T. A. Royer, S. Biles, E. Maxson, Sugarcane Aphid (Hemiptera: Aphididae): A New Pest on Sorghum in North America. *J. Integr. Pest Manag.* **7** (2016), doi:10.1093/jipm/pmw011.
17. J. W. Gordy, M. J. Brewer, R. D. Bowling, G. D. Buntin, N. J. Seiter, D. L. Kerns, F. P. F. Reay-Jones, M. O. Way, Development of Economic Thresholds for Sugarcane Aphid (Hemiptera: Aphididae) in Susceptible Grain Sorghum Hybrids. *J. Econ. Entomol.* **112**, 1251–1259 (2019).

18. S. Nibouche, L. Costet, J. R. Holt, A. Jacobson, A. Pekarcik, J. Sadeyen, J. S. Armstrong, G. C. Peterson, N. McLaren, R. F. Medina, Invasion of sorghum in the Americas by a new sugarcane aphid (Melanaphis sacchari) superclone. *PLOS ONE*. **13**, e0196124 (2018).
19. USDA-FAS, Grain and Feed Annual Report: Haiti (HA2020-0001) (2020), (available at https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Grain%20and%20Feed%20Annual_Port-au-Prince_Haiti_04-15-2020).
20. W. Calvin, Effects of intercropping and biological insecticides on sugarcane aphid (Hemiptera: aphididae) infestations on sorghum, and identification of natural enemies and alternate hosts in Haiti, 77 (2019).
21. K. T. Muleta, G. Pressoir, G. P. Morris, Optimizing genomic selection for a sorghum breeding program in Haiti: a simulation study. *G3 Genes Genomes Genet.* **9**, 391–401 (2019).
22. F. Wang, S. Zhao, Y. Han, Y. Shao, Z. Dong, Y. Gao, K. Zhang, X. Liu, D. Li, J. Chang, D. Wang, Efficient and fine mapping of RMES1 conferring resistance to sorghum aphid Melanaphis sacchari. *Mol. Breed.* **31**, 777–784 (2013).
23. J. S. Armstrong, W. L. Rooney, G. C. Peterson, R. T. Villenueva, M. J. Brewer, D. Sekula-Ortiz, Sugarcane Aphid (Hemiptera: Aphididae): Host Range and Sorghum Resistance Including Cross-Resistance From Greenbug Sources. *J. Econ. Entomol.* **108**, 576–582 (2015).
24. J. S. Armstrong, L. Mbulwe, D. Sekula-Ortiz, R. T. Villanueva, W. L. Rooney, Resistance to Melanaphis sacchari (Hemiptera: Aphididae) in forage and grain sorghums. *J. Econ. Entomol.* **110**, 259–265 (2017).
25. C. Burgarella, A. Barnaud, N. A. Kane, F. Jankowski, N. Scarcelli, C. Billot, Y. Vigouroux, C. Berthouly-Salazar, Adaptive Introgression: An Untapped Evolutionary Mechanism for Crop Adaptation. *Front. Plant Sci.* **10** (2019), doi:10.3389/fpls.2019.00004.
26. H. M. Tetreault, S. Grover, E. D. Scully, T. Gries, N. A. Palmer, G. Sarath, J. Louis, S. E. Sattler, Global Responses of Resistant and Susceptible Sorghum (Sorghum bicolor) to Sugarcane Aphid (Melanaphis sacchari). *Front. Plant Sci.* **10** (2019), doi:10.3389/fpls.2019.00145.
27. Y. Jiang, C.-X. Zhang, R. Chen, S. Y. He, Challenging battles of plants with phloem-feeding insects and prokaryotic pathogens. *Proc. Natl. Acad. Sci.* **116**, 23390–23397 (2019).
28. J. R. Andersen, T. Lübberstedt, Functional markers in plants. *Trends Plant Sci.* **8**, 554–560 (2003).
29. L. McHale, X. Tan, P. Koehl, R. W. Michelmore, Plant NBS-LRR proteins: adaptable guards. *Genome Biol.* **7**, 212 (2006).
30. C. W. Smith, R. A. Frederiksen, in *Sorghum: Origin, History, Technology, and Production*, C. W. Smith, R. A. Frederiksen, Eds. (Wiley, NY, 2000; https://books.google.com/books?hl=en&lr=&id=b7vxU44v794C&oi=fnd&pg=PA191&dq=History+of+Cultivar+Development+in+the+United+States:+From+%22Memoirs+of+A.+B.+Maunder--Sorghum+Breeder%22&ots=UuB7XqgYWP&sig=cVMz_e6MAD5WTEKprcSdYOThAs0), vol. 2, p. 191.
31. R. R. Klein, F. R. Miller, D. V. Dugas, P. J. Brown, A. M. Burrell, P. E. Klein, Allelic variants in the PRR37 gene and the human-mediated dispersal and diversification of sorghum. *Theor. Appl. Genet.*, 1–15 (2015).

32. G. C. Peterson, J. W. Johnson, G. L. Teetes, D. T. Rosenow, Registration of Tx2783 Greenbug Resistant Sorghum Germplasm Line. *Crop Sci.* **24**, 390 (1984).
33. République du Sénégal, Catalogue officiel des espèces et des variétés cultivées au Sénégal [Official catalog of cultivated species and varieties in Senegal] (2012), (available at http://www.fao.org/pgrfa-gpa-archive/sen/docs/senegal_varietes/Catalogue_%20varietes.htm).
34. J. M. Faye, F. Maina, E. A. Akata, B. Sine, C. Diatta, A. Mamadou, S. Marla, S. Bouchet, N. Teme, J.-F. Rami, D. Fonceka, N. Cisse, G. P. Morris, A genomics resource for genetics, physiology, and breeding of West African sorghum. *Plant Genome*. **n/a**, e20075 (2021).
35. L. Mbulwe, G. C. Peterson, J. Scott-Armstrong, W. L. Rooney, Registration of sorghum germplasm Tx3408 and Tx3409 with tolerance to sugarcane aphid [Melanaphis sacchari (Zehntner)]. *J. Plant Regist.* **10**, 51–56 (2016).
36. G. C. Peterson, J. S. Armstrong, B. B. Pendleton, M. Stelter, M. J. Brewer, Registration of RTx3410 through RTx3428 sorghum germplasm resistant to sugarcane aphid [Melanaphis sacchari (Zehntner)]. *J. Plant Regist.* **12**, 391–398 (2018).
37. J. Ndjeunga, K. Mausch, F. Simtowe, in *Crop improvement, adoption, and impact of improved varieties in food crops in sub-Saharan Africa*, T. S. Walker, J. Alwang, Eds. (CABI, Wallingford, 2015; http://www.cabi.org/cabebooks/ebook/20153367543), pp. 123–147.
38. S. V. Nuzhdin, T. L. Turner, Promises and limitations of hitchhiking mapping. *Curr. Opin. Genet. Dev.* **23** (2013), doi:10.1016/j.gde.2013.10.002.
39. S. Hoban, J. L. Kelley, K. E. Lotterhos, M. F. Antolin, G. Bradburd, D. B. Lowry, M. L. Poss, L. K. Reed, A. Storfer, M. C. Whitlock, Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *Am. Nat.* (2016), doi:10.1086/688018.
40. T. Yohannes, T. Abraha, D. Kiambi, R. Folkertsma, C. T. Hash, K. Ngugi, E. Mutitu, N. Abraha, M. Weldetsion, C. Mugoya, C. W. Masiga, S. de Villiers, Marker-assisted introgression improves Striga resistance in an Eritrean farmer-preferred sorghum variety. *Field Crops Res.* **173**, 22–29 (2015).
41. T. A. Royer, B. B. Pendleton, N. C. Elliott, K. L. Giles, Greenbug (hemiptera: aphididae) biology, ecology, and management in wheat and sorghum. *J. Integr. Pest Manag.* **6** (2015), doi:10.1093/jipm/pmv018.
42. F. R. Miller, Y. Kebede, in *Genetic Contributions to Yield Gains of Five Major Crop Plants* (John Wiley & Sons, Ltd, 1984; https://acsess.onlinelibrary.wiley.com/doi/abs/10.2135/cssaspecpub7.c1), pp. 1–14.
43. M. J. Thomson, High-throughput SNP genotyping to accelerate crop improvement. *Plant Breed. Biotechnol.* **2**, 195–212 (2014).
44. P. Heisey, K. Day-Rubenstein, "Using crop genetic resources to help agriculture adapt to climate change: economics and policy" (SSRN Scholarly Paper ID 2709190, Social Science Research Network, Rochester, NY, 2015), , doi:10.2139/ssrn.2709190.
45. J. Kloppenburg, D. L. Kleinman, The plant germplasm controversy. *BioScience*. **37**, 190–198 (1987).
46. J. Kotschi, B. Horneburg, The Open Source Seed Licence: A novel approach to safeguarding access to plant germplasm. *PLOS Biol.* **16**, e3000023 (2018).
47. R. Bernardo, *Molecular markers and selection for complex traits in plants: Learning from the last 20 years, Crop Science 48 (2008): 1649-1664.* (2008), vol. 48.

48. B. Brachi, G. P. Morris, J. O. Borevitz, Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biol.* **12**, 232 (2011).
49. FAO, FAOSTAT (2021), (available at http://www.fao.org/faostat/en/).
50. G. P. Morris, P. Ramu, S. P. Deshpande, C. T. Hash, T. Shah, H. D. Upadhyaya, O. Riera-Lizarazu, P. J. Brown, C. B. Acharya, S. E. Mitchell, J. Harriman, J. C. Glaubitz, E. S. Buckler, S. Kresovich, Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc. Natl. Acad. Sci.* **110**, 453–458 (2013).
51. J. R. Lasky, H. D. Upadhyaya, P. Ramu, S. Deshpande, C. T. Hash, J. Bonnette, T. E. Juenger, K. Hyma, C. Acharya, S. E. Mitchell, E. S. Buckler, Z. Brenton, S. Kresovich, G. P. Morris, Genome-environment associations in sorghum landraces predict adaptive traits. *Sci. Adv.* **1**, e1400218 (2015).
52. J. C. Glaubitz, T. M. Casstevens, F. Lu, J. Harriman, R. J. Elshire, Q. Sun, E. S. Buckler, TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*. **9**, e90346 (2014).
53. R. F. McCormick, S. K. Truong, A. Sreedasyam, J. Jenkins, S. Shu, D. Sims, M. Kennedy, M. Amirebrahimi, B. D. Weers, B. McKinley, A. Mattison, D. T. Morishige, J. Grimwood, J. Schmutz, J. E. Mullet, The Sorghum bicolor reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *Plant J.* **93**, 338–354 (2018).
54. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. **25**, 1754–1760 (2009).
55. B. L. Browning, S. R. Browning, Genotype imputation with millions of reference samples. *Am. J. Hum. Genet.* **98**, 116–126 (2016).
56. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics*. **27**, 2156–2158 (2011).
57. T. Jombart, S. Devillard, F. Balloux, Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* **11**, 94 (2010).
58. P. J. Bradbury, Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss, E. S. Buckler, TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*. **23**, 2633–2635 (2007).
59. E. Paradis, J. Claude, K. Strimmer, APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*. **20**, 289–290 (2004).
60. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
61. M. C. Whitlock, K. E. Lotterhos, Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of FST. *Am. Nat.* **186**, S24–S36 (2015).
62. G. A. V. der Auwera, M. O. Carneiro, C. Hartl, R. Poplin, G. del Angel, A. Levy□Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S. Gabriel, M. A. DePristo, *Curr. Protoc. Bioinforma.*, in press, doi:https://doi.org/10.1002/0471250953.bi1110s43.
63. J. S. Armstrong, W. L. Rooney, G. C. Peterson, R. T. Villenueva, M. J. Brewer, D. Sekula-Ortiz, Sugarcane aphid (Hemiptera: Aphididae): host range and sorghum resistance including cross-resistance from greenbug sources. *J. Econ. Entomol.* **108**, 576–582 (2015).

64. S. Paudyal, J. S. Armstrong, K. L. Giles, M. E. Payton, G. P. Opit, A. Limaje, Categories of resistance to sugarcane aphid (hemiptera: aphididae) among sorghum genotypes. *J. Econ. Entomol.* **112**, 1932–1940 (2019).